BELLCOMM, INC.
955 L'ENFANT PLAZA NORTH, S.W. WASHINGTON, D.C. 20024

COVER SHEET FOR TECHNICAL MEMORANDUM

TITLE- The Simulation of a Stationary Gaussian Stochastic Process with Rational Spectral Density

TM- 68-1033-7

DATE- December 31, 1968

FILING CASE NO(S)- 620
101

AUTHOR(S)- J. L. Strand

FILING SUBJECT(S)- Simulation of Stochastic Processes
(ASSIGNED BY AUTHOR(S)- Stationary Processes

ABSTRACT

In this report we give explicit techniques for generating on a digital computer sample paths of any stationary Gaussian process with a rational spectral density (or, equivalently, of any stationary Gaussian Markov process). This is done by reducing the problem to one of generating solutions of a linear differential equation driven by white noise. Then an explicit method for solving the second problem is given.

Several methods of computing the statistical properties of any stationary solution of any linear differential equation driven by a stationary process are developed. These are then applied to our particular problem.

The report ends with a study of two alternative methods of testing the generated sample paths to make sure that they do indeed have the desired statistical properties.

SEE REVERSE SIDE FOR DISTRIBUTION LIST

BELLCOMM, INC. TM- 68-1033-7

DISTRIBUTION

COMPLETE MEMORANDUM TO

CORRESPONDENCE FILES:

OFFICIAL FILE COPY
plus one white copy for each
additional case referenced

TECHNICAL LIBRARY (4)

Bellcomm

Messrs. G. R. Andersen
G. M. Anderson
A. P. Boysen, Jr.
K. R. Carpenter
D. A. Chisholm
B. D. Elrod
J. J. Fearnsides
D. R. Hagner
H. A. Helm
B. T. Howard
J. Kranton
W. Levidow
J. Z. Menard
J. M. Nervik
I. M. Ross
J. W. Schindelin
P. G. Smith
J. W. Timko
R. L. Wagner
Mrs. N. I. Kirkendall
Mrs. L. L. Wang
All Members of Department 1033
Department 1024 File

COVER SHEET ONLY TO

BA-146 (8-64)

BELLCOMM, INC.
955 L'ENFANT PLAZA NORTH, S.W. WASHINGTON, D.C. 20024

SUBJECT: The Simulation of a Stationary Gaussian Stochastic Process with Rational Spectral Density - Cases 620, 101

DATE: December 31, 1968

FROM: J. L. Strand

TM-68-1033-7

TECHNICAL MEMORANDUM

1. INTRODUCTION

Simulation problems often require a random input with specified statistical properties. This report develops a method for generating such an input in one important case, namely when the input is to be a stationary Gaussian process with a rational spectral density (or equivalently, when it is to be a stationary Gaussian Markov process). The method explicitly shows how to simulate precisely any such stationary process on a digital computer and how to test the generated data to make sure that the output has the desired statistical properties.

In section two we show that the problem of generating sample paths of the desired process is equivalent to the problem of solving a certain stochastic differential equation. In section three we solve this equation and compute its statistical properties. In section four techniques for generating sample paths of the desired process on a digital computer are discussed. Section five develops two ways of testing the simulation.

Two appendices give mathematical justification of some of the techniques used.

Throughout this paper the following notation will be used: y(t) will be the stationary Gaussian process we are studying. We assume y has zero mean. If A is a random variable, EA will be its mean.

$$B(h) = Ey(t)y(t+h) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} y(t)y(t+h)dt$$

is the covariance (autocorrelation) of y. $f(\lambda)$ is the spectral density of B (or the "power spectral density" of y). Then

$$(1)\qquad \begin{cases} B(h) = \displaystyle\int_{-\infty}^{\infty} e^{j\lambda h} f(\lambda)\,d\lambda \\ f(j\lambda) = \dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty} e^{-j\lambda h} B(h)\,dh \end{cases} .$$

We assume $f(\lambda)$ is rational. Then since B is an even real function so is f. Consequently

$$(2)\qquad f(j\lambda) = \frac{|P(j\lambda)|^2}{|Q(j\lambda)|^2} = \left|\sum_{i=0}^{m} b_i (j\lambda)^i\right|^2 \Bigg/ \left|\sum_{i=0}^{n} a_i (j\lambda)^i\right|^2$$

where $m<n$, and P and Q have only roots with negative real parts.

Underlined capital letters are always nxn matrices. Underlined lower case letters are column n-vectors and $\underline{x} = \text{col}(x_1, \ldots, x_n)$. $\underline{I}$ is the nxn identity matrix.

If $\mathbf{x}$ and $\underline{y}$ are random column n-vectors, $\text{cov}(\underline{x},\underline{y}) = E(\underline{x}\underline{y}^T)$.

2. DERIVATION OF THE DIFFERENTIAL EQUATION

$y(t)$ is the stationary Gaussian process (with statistical properties defined by (1) and (2)) that we wish to simulate. In this section we prove the following result:

$$y(t) = \sum_{k=0}^{m} b_k x^{(k)}(t)$$

where $x(t)$ is the "unique stationary solution" of the differential equation

$$(3)\qquad a_n x^{(n)}(t) + \ldots + a_0 x(t) = u(t) \quad ,$$

where u is white noise with zero mean and covariance $\delta(h)$. (By "unique stationary solution" we mean the following: the initial conditions $x(0)$, $x^1(0)$, ..., $x^{(n-1)}(0)$ must be chosen from a unique n-variate normal distribution in order that x(t) be stationary. In this case the statistical properties of x are uniquely determined.)

Equation (3) may be written in matrix form: let $x_1(t) = x(t)$, $x_2(t) = x^1(t)$, ..., $x_n(t) = x^{(n-1)}(t)$. Then $x_k'(t) = x_{k+1}(t)$ if $k<n$, and

$$x_n'(t) = -(a_0x_1(t) + a_1x_2(t) + \dots + a_{n-1}x_n(t))/a_n + u(t)/a_n$$

If we let $\underline{x}(t) = \mathrm{col}(x_1(t),\dots,x_n(t))$ and

$$\underline{A} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \\ 0 & \dots & \dots & 0 & 1 \\ -a_0/a_n & & & & -a_{n-1}/a_n \end{pmatrix}$$

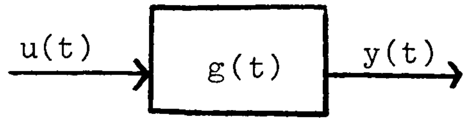
Then (3) becomes

(4) $$\underline{x}'(t) = \underline{Ax}(t) + \underline{w}(t)$$

where $\underline{w}(t) = \mathrm{col}\ (0, \dots, 0, u(t)/a_n)$.

We now give an intuitive proof of the assertion given at the start of this section. A rigorous proof of the same assertion is given in Appendix 1.

Consider a linear system with causal impulse response g(t) driven by white noise u(t):



Assuming that the system has achieved its steady state, it satisfies

$$(5) \qquad y(t) = \int_{-\infty}^{\infty} g(\alpha)u(t-\alpha)d\alpha = \int_{-\infty}^{t} g(t-\alpha)u(\alpha)d\alpha \quad .$$

Clearly $Ey(t) = 0$. The covariance of y is easily computed:

$$Ey(t+h)y(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(\alpha_1)g(\alpha_2)\delta(h+\alpha_2-\alpha_1)d\alpha_1 d\alpha_2$$

and so y has covariance

$$Ey(t+h)y(t) = \int_{-\infty}^{\infty} g(\alpha+h)g(\alpha)d\alpha \quad .$$

Taking Fourier transforms we find that y has power spectral density $|G(j\omega)|^2$ where $G(j\lambda) = \int_{-\infty}^{\infty} g(t)e^{-j\lambda t}dt$. Hence if we let $G(j\lambda) = P(j\lambda)/Q(j\lambda)$ the linear system (5) with $g(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} G(j\lambda)e^{j\lambda t}d\lambda$ has an output process $y(t)$ with the desired power spectral density (2).

Returning to (5) and taking Fourier transforms

$$(\hat{u}(j\lambda) = \int_{-\infty}^{\infty} e^{-j\lambda t}u(t)dt, \quad \hat{y}(j\lambda) = \int_{-\infty}^{\infty} e^{-j\lambda t}y(t)dt)$$

we see $\hat{y}(j\lambda) = \hat{u}(j\lambda)P(j\lambda)/Q(j\lambda)$. Using equation (2) we see that

$$\sum_{k=0}^{n} a_k (j\lambda)^{(k-n)} \hat{y}(y\lambda) = \sum_{k=0}^{n} b_k (j\lambda)^{(k-n)} \hat{u}(j\lambda) \quad .$$

Taking inverse transforms this give

$$\sum_{0}^{n} a_k y^{(k-n)}(t) = \sum_{0}^{m} b_k u^{(k-n)}(t) \quad . \tag{6}$$

Therefore if x(t) is a stationary solution of $\sum_{k=0}^{n} a_k{}^{(k)}(t) = u(t)$, clearly $y(t) = \sum b_k x^{(k)}(t)$ solves (6), as desired.

A mathematically rigorous proof of this result is given in Appendix 1. From a mathematical point of view a rigorous alternative proof is desirable for the following reasons:

1. White noise has mathematical meaning only as the limit of a sequence of Gaussian processes with decreasing correlation times. Hence equation (5), for example, is mathematically meaningless unless a more advanced approach is used.

2. Whether or not the linear system in question has a unique steady state when driven by white noise is not mathematically clear.

3. Taking Fourier transforms of (5) is not in general possible. Hence the replacement of (5) by (6) needs more justification.

3. STATIONARY SOLUTIONS OF RANDOM DIFFERENTIAL EQUATIONS

In the last section we reduced our simulation problem to the problem of simulating the stationary solution of the linear stochastic equation (4) $\underline{x}'(t) = \underline{Ax}(t) + \underline{w}(t)$ where A has only eigenvalues with negative real parts. In this section we compute the statistical properties of the stationary solution of a slightly more general equation, namely

$$\underline{x}'(t) = \underline{Bx}(t) + \underline{R}(t) \tag{7}$$

where $\underline{R}(t) = \mathrm{Col}(R_1,\ldots,R_n)$ is some stationary process with zero mean and $\underline{B}$ is a matrix all of whose eigenvalues have negative real parts. The results are then specialized to our case.

Let $\underline{\Phi}(t)$ be the matrix solution of $\underline{\Phi}'(t) = B\underline{\Phi}(t)$, $\underline{\Phi}(0) = I$. The following properties of $\underline{\Phi}$ are well known (see any book on ordinary differential equations, for example [2]).

a. $\underline{\xi}(t) = \mathrm{col}(\Phi_{1k}(t),\ldots,\Phi_{nk}(t))$ solves $\underline{\xi}' = \underline{B\xi}$, $\underline{\xi}(0) = (0,\ldots,1,\ldots,0)$ (the 1 in k-th position).

b. If $\underline{x}(t)$ solves (7) then

$$\underline{x}(t) = \underline{\Phi}(t-a)\underline{x}(a) + \int_a^t \underline{\Phi}(t-s)\underline{R}(s)ds \quad . \tag{8}$$

We now show that not only does (7) have a unique stationary solution but $\underline{x}$ is a stationary solution of (7) if and only if

$$\underline{x}(t) = \int_{-\infty}^t \underline{\Phi}(t-s)\underline{R}(s)ds \quad . \tag{9}$$

This is a consequence of (8): If $\underline{x}$ is a stationary solution then Var $(\underline{x}(s))$ is a constant, and since $|\Phi(t-a)| \to 0$

as $a \to -\infty$ we get (9). Conversely, as defined by (9) x(t) is stationary if R(s) is and is even strictly stationary if R is Gaussian. In fact

$$E\underline{x}(t) = \int_{-\infty}^{t} \underline{\Phi}(t-s)E\underline{R}(s)ds = 0$$

$$(10)\quad \mathrm{cov}(\underline{x}(t_1),\underline{x}(t_2)) = \int_{-\infty}^{t_1}\int_{-\infty}^{t_2} \Phi(t_1-s_1)\Gamma(s_1-s_2)\Phi^T(t_2-s_2)ds_1ds_2$$

where $\Gamma(t) = E(\underline{R}(s)\underline{R}^T(s+t))$. This clearly is a function of t_1-t_2 only. Hence $\underline{x}$ is (second order) stationary. If $\underline{R}$ is Gaussian so is $\underline{x}$ and in this case second-order stationarity implies strict stationarity.

If Γ or Φ are complicated functions the integral in (10) may be impossible to compute. For example, if the matrix B is very large, computing Φ is very hard. In the important special case where our matrix equation (7) arises from a linear n-th order equation

$$(11)\qquad x^{(n)}(t) + b_{n-1}x^{(n-1)}(t) \ldots + b_0x(t) = R(t)$$

with stationary driving function R(t) Karhunen [3] gives an alternative method which may be feasible when the computation of (10) is not feasible. This method is described in Appendix 2.

We now specialize (10) to the special case (4) where the driving function is white noise. In this case (9) becomes $\underline{x}(t) = \int_{-\infty}^{t} \underline{\Phi}(t-s)\underline{w}(s)ds$. $D = E(\underline{w}(0)\underline{w}^T(t))$ is zero except for $d_{nn} = \delta(t)/a_n^2$. Hence we have the following complete characterization of the solution $\underline{x}$ of (4).

(1) $E\underline{x}(t) = 0$

(2) $$\text{cov}(\underline{x}(t+h)\underline{x}^T(t)) = E\int_{-\infty}^{t+h}\int_{-\infty}^{t} \underline{\Phi}(t+h-s_1)\underline{w}(s_1)\underline{w}^T(s_2)\underline{}^T(t-s)ds_1 ds_2$$

$$= \int_{-\infty}^{t} \underline{\Phi}(t+h-s)\underline{D\Phi}^T(t-s)ds$$

Hence

(13) $$\text{cov}(\underline{x}(t+h)\underline{x}^T(t)) = \int_0^{\infty} \underline{\Phi}(s+h)\underline{D\Phi}^T(s)ds$$

or (component-wise)

$$E(x_i(t+h)x_j(t)) = \left(\int_0^{\infty} \phi_i(s+h)\phi_j(s)ds\right) \Big/ a_n^2$$

where $\underline{\Phi} = (\phi_{ij})$.

(3) $\underline{x}$ is Gaussian.

4. DIGITAL APPROXIMATION OF x(t)

For a given step size T we wish to construct sample paths $\underline{x}(nT)$, n = 0, 1, ..., N with statistical properties agreeing with those of $\underline{x}$ in equation (13). We start by choosing $\underline{x}(0)$ from an n-variate Gaussian distribution with zero mean and variance (from equation (13)) $\text{Var}(\underline{x}(0)) = \int_0^{\infty} \underline{\Phi}(\nu)\underline{D\Phi}^T(\nu)d\nu$.

By equation (8),

$$\underline{x}((n+1)T) = \underline{\Phi}(T)\underline{x}(nT) + \int_{nT}^{(n+1)T} \underline{\Phi}((n+1)T-s)\underline{u}(s)ds$$

or

(14) $$\underline{x}((n+1)T) = \underline{\Phi}(T)\underline{x}(nT) + \underline{w}_{n+1}$$

where $\underline{w}_{n+1}$ is a Normal random variable, independent of $\underline{x}(nT)$, with mean 0 and variance

$$\int_0^T \underline{\Phi}(s)\underline{B\Phi}^T(s)ds \quad .$$

If the state vector is large, the computation of Φ is difficult. However, if it is possible somehow to find the initial distribution cov($\underline{x}(0)$, $\underline{x}(0)$) we point out that one could then do a straighforward numerical integration of (7), or use the Runge Kutta method of solving ordinary differential equations numerically. In this case the driving function R(t) is replaced by a sequence $\{\underline{R}(nT)\}$ of independent normal random variables.

One case where the initial values can be computed directly even for large n is the n-th order linear case discussed in Appendix 2.

We conclude this section by outlining a method for generating an n-dimension normal random vector $\underline{x}$ with a given covariance matrix Γ. This is done in two steps:

(1) Generate an n-dimensional random vector $\underline{n}$ with independent components so $E(\underline{nn}^T) = \underline{I}$.

(2) Make a linear transformation $\underline{x} = \underline{C}n$ with an appropriate deterministic matrix $\underline{C}$. To discover the proper form for $\underline{C}$ we reason as follows: if $\underline{x} = \underline{Cn}$. then $E(\underline{xx}^T) = E(\underline{Cnn}^T C^T) = \underline{CC}$. Hence $\underline{C}$ need only satisfy $\underline{CC}^T = \Gamma$.

$\underline{C}$ may be taken to be a lower-triangular matrix and constructed as follows: Let $\Gamma = CC^T$, $\Gamma = (\gamma_{ij})$, $C = (C_{ij})$. Since $C_{ij} = 0$ if $i < j$, it is easily seen that the C_{ij} can be found recursively as follows:

$$\gamma_{11} = C_{11}^2 \quad , \qquad C_{11} = \sqrt{\gamma_{11}}$$

$$\gamma_{1k} = a_{11}a_{k1} \qquad C_{k1} = \gamma_{1k}/a_{11}$$

and in general the i-th row of $\underline{C}$ can be determined from the equation

$$\gamma_{ij} = \sum_{k=1}^{\mathrm{Min}(i,j)} C_{ik}\, C_{jk}$$

which can be solved for C_{ij} successively.

5. TESTING THE DIGITAL SIMULATION

We wish to test the sequence $\{\underline{x}(rT)\}$ generated by equation (14) to see if its statistical behavior agrees with that calculated theoretically in section 3. In essence we are just testing the random number generator used to generate our "white noise" sequence $\underline{w}_r$ - if we had a perfect random number generator, our sequence $\underline{x}(rT)$ would have precisely the desired statistical properties. Since this will not be, our sequence will vary from that desired.

We suppose that M sample paths of $\underline{x}$ are available, $\underline{x}^1(rT)$, $\underline{x}^2(rT)$, ..., $\underline{x}^M(rT)$ where r=0, 1,..., N and $\underline{x}^k(rT) = \mathrm{col}(x_1^k(rT), \ldots, x_n^k(rT))$. We assume that these sample paths are truly independent. Define $p_{ij}(kT) = E(x_i(0)x_j(kT))$, and let $\underline{p}(kT) = (p_{ij}(kT))_{i,j=1}^n$.

There are four statistical properties of $\underline{x}$ we would like to test--stationarity, normality, proper mean, proper covariance. There seem to be no useable tests of stationarity, but this can be indirectly checked through the mean and covariance functions. Testing for normality is a very well known statistical problem. We refer the reader to [4], for example, for information.

There are two alternate sets of statistics one can use to test the mean and covariance of $\underline{x}$. The first of these is

$$(15)\qquad \begin{cases} \overline{\underline{x}}(rT) = \frac{1}{M} \sum_{k=1}^{M} \underline{x}^k(rT) \\ \\ \overline{p}_{ij}(rT,(r+\ell)T = \frac{1}{M} \sum_{k=1}^{M} x_i^k(rT)x_j^k((r+\ell)T) \end{cases}$$

We let $\overline{\underline{p}}(rT,(r+\ell)T)$ be the matrix $(\overline{p}_{ij}(rT,(r+\ell)T))$.

The second set of statistics is

$$(16)\qquad \begin{cases} \overline{\underline{x}}^k = \frac{1}{N+1} \sum_{\ell=0}^{N} \underline{x}^k(\ell T) \\ \\ \beta_{ij}^k(rT) = \frac{1}{N+1-r} \sum_{\ell=0}^{N-r} x_i^k(\ell T)x_j^k((\ell+r)T) \quad . \end{cases}$$

Let $\underline{\beta}^k(rT) = (\beta_{ij}^k(rT))_{i,j=1}^n$.

An evaluation of these two approaches is given later. First we compute the means and variances of the statistics.

Clearly $E(\overline{\underline{x}}(rT)) = \frac{1}{M} \sum_k E\underline{x}(rT) = 0$. We calculate the variance of $\overline{\underline{x}}(rT)$ component by component:

$$(17)\quad E(\overline{x}_i(rT)\overline{x}_j(rT)) = \frac{1}{M^2} \sum_{k=1}^{M} \sum_{\ell=1}^{M} E(x_i^k(rT)x_j^\ell(rT)) = \frac{p_{ij}(0)}{M} \ .$$

Hence $\mathrm{Var}(\overline{\underline{x}}(rT)) = \frac{1}{M}\, \underline{p}(0)$.

The expectation of $\overline{\underline{p}}(rT,(r+\ell)T) = \underline{p}(\ell T)$ obviously. The variance is given component by component:

$$(18)\quad E(\overline{p}_{ij}(rT,(r+\ell)T) - p_{ij}(\ell T))^2 = \frac{1}{M}[p_{ii}(0)p_{jj}(0) + p_{ij}^2(0)] \quad .$$

To prove this we compute

$$(\overline{p}_{ij}(rT,(r+\ell)T))^2 = \frac{1}{M^2}\sum_{k_1}\sum_{k_2} x_i^{k_1}(rT)x_j^{k_1}((r+\ell)T)x_i^{k_2}(rT)x_j^{k_2}((r+\ell)T)$$

$$+ \frac{1}{M^2}\sum_{k=1}^{M} (x_i^k(rT)x_j^k((r+\ell)T))^2$$

$$+ \sum_{k_1=2}^{M}\sum_{k_1=1}^{M} x_i^{k_1}(rT)x_j^{k_1}((r+\ell)T)x_i^{k_2}(rT)x_j^{k_2}((r+\ell)T)$$

To compute the expectation of these terms we use a characteristic function argument. If $v_1, \ldots, v_n$ are any random variables and $(t_1,\ldots,t_n) = Ee^{i[t_1v_1+\ldots+t_nv_n]}$ and $k=k_1+k_2+\ldots+k_n$ then $(i)^kE(x_1^{k_1}\ldots x_n^{k_n}) = \dfrac{\partial^k}{\partial t_1^{k_1}\ldots\partial t_n^{k_n}}\,\phi(0,0,\ldots,0).$ Applying this

$$E(\overline{p}_{ij}(rT,(r+\ell)T))^2 = \frac{1}{M^2}[M(p_{ii}(0)p_{jj}(0)+2p_{ij}^2(\ell T)) + M(M-1)(p_{ij}^2(\ell T))]$$

$$= \frac{1}{M}[p_{ii}(0)p_{jj}(0) + p_{ij}^2(\ell T)] + p_{ij}^2(\ell T)$$

from which (18) follows immediately.

We now turn to the statistics (16). Clearly $E(\underline{\bar{x}}^k) = 0$, $E(\beta_{ij}^k(rT)) = p_{ij}(rT)$. Hence the statistics are unbiased. $\mathrm{Cov}(\bar{x}^k)$ will be computed term by term.

$$E(\bar{x}_i^k \bar{x}_j^k) = \frac{1}{(N+1)^2} E\left[\sum_{\ell=0}^{N} (x_i^k(\ T)x_j^k(\ T) + 2 \sum_{\ell_1=1}^{N} \sum_{\ell_2=0}^{\ell_1-1} x_i^k(\ell_1 T)x_j^k(\ell_2 T)\right]$$

$$= \frac{1}{(N+1)^2} \left[(N+1)p_{ij}(0) + 2 \sum_{\ell=1}^{N} (N+1-\ell)p_{ij}(\ell T)\right] .$$

So

$$(19) \qquad E(\bar{x}_i^k \bar{x}_j^k) = \frac{1}{N+1} \left[p_{ij}(0) + 2 \sum_{\ell=1}^{N} (1-\frac{\ell}{N+1})\ p_{ij}(\ell T)\right] .$$

To compute the variance of $\beta_{ij}^k\ (rT)$, note that if $z(t) = x_i^k(t)x_j^k(t+rT) - p_{ij}(rT) - p_{ij}(rT)$, $x(t)$ is stationary with $Ez = 0$ and

$$\mathrm{Var}(\beta_{ij}^k(rT)) = (\frac{1}{N+1-r})^2 \sum_{\ell_1=0}^{N-r} \sum_{\ell_2=0}^{N-r} E(z(\ell_1 T)z(\ell_2 T)) .$$

But an argument exactly like those used above then shows that

$$\mathrm{Var}(\beta_{ij}^k(rT)) = \frac{1}{N+1-r} \left[E(z^2(0)) + 2 \sum_{\ell=1}^{N-r}(r-\frac{\ell}{1+N-r})E(z(0)z(\ell T))\right]$$

But it is easily seen (again by using characteristic) functions) that $E(z(0)z(\ell T) = p_{ii}(\ell T)p_{jj}(\ell T) + p_{ij}((\ell+r)T)p_{ij}((\ell-r)T)$ and we have

$$(20) \quad Var(\beta^k_{ij}(rT)) = \frac{1}{N+1-r}\Bigg[p_{ii}(0)p_{jj}(0) + p^2_{ij}(rT)$$
$$+ \sum_{\ell=1}^{N-r}(1-\frac{\ell}{N+1-r})(p_{ii}(\ell T)p_{jj}(\ell T)$$
$$+ p_{ij}((\ell+r)T)p_{ij}((\ell-r)T))\Bigg]$$

In order to use the statistics (15), the following steps should be followed:

1. Each time a new sample path $\underline{x}^k$ is generated the number $\{\underline{x}^k(rT)\}$ must be preserved.

2. When a sufficient number of sample paths have been computed, form the statistics $\overline{\underline{x}}(rT)$ and $\overline{\rho}_{ij}(rT)$, $r=0, 1, \ldots, N$.

3. $\overline{\underline{x}}(rT)$ should have mean zero and variance given by (17). Being the sum of independent normal random variables $\overline{\underline{x}}(rT)$ should be normal. Hence a comparison of the observed $\overline{\underline{x}}(rT)$ and the standard deviation is a test of system performance. Similarly, ρ_{ij} will be approximately normal for large M by the central limit theorem, and hence a standard deviation test can be applied to it, using (18).

In order to use the statistics (16), the following steps should be followed:

1. During the computation of each sample path $\underline{x}^k$ the statistics should be cumulated, and stored at the end of each run.

2. After each run, a standard deviation test on $\overline{x}^k$ can be performed; $\overline{x}^k$ should be normally distributed with mean and variance (19). Unfortunately, the distribution of β_{ij} is not known and hence the only test that can be made each sample run would be to use Chebyshev's inequality. This is notorously pessimistic and will give useable information only if N is very large. However,

3. After a number of sample paths have been created we can form the statistics

$$\beta_{ij}(rT) = \frac{1}{M} \sum_{k=1}^{M} \beta_{ij}^{k}(rT)$$

Since the $\overline{\beta}^k$ are independent $\overline{\beta}(rT)$ should be normally distributed with mean $\rho_{ij}(rT)$ and variance $\dfrac{\mathrm{Var}(\beta_{ij}^{k}(rT))}{M}$.

Each of the statistics (15) and (16) have some advantages. The distributions of the statistics (15) are more easily computed, especially when compared to $\overline{\rho}_{ij}^{k}(rT)$. However, many more numbers need to be preserved than are needed in (16)--essentially N^2+N versus 2N. Also the programming involved to compute (16) should be much less than that required to compute (15). A second factor to consider is the type of assumptions underlying our analysis of each set of statistics. The statistics (15) may be analyzed essentially assuming only sample path independence. The analysis of the statistics (16) assumes in addition the stationarity of the process.

Based upon this analysis, it is recommended that

1. Primary reliance be put on the statistics (16). However, to check the stationarity assumption a few values of $\overline{\rho}_{ij}$ should be computed, especially the cases r=0, ℓ=0 and ℓ=N.

2. After initial data has been analyzed a re-evaluation of the statistical techniques being used should be made.

J. L. Strand

1033-JLS-jr

Attachments
Appendix 1, 2
References

APPENDIX 1

DERIVATION OF THE DIFFERENTIAL EQUATION

In this appendix we give a proof of the result stated at the start of section 2: there is a unique probability distribution from which the initial conditions for equation (3) may be chosen which leads to a stationary solution of (3). If $x(t)$ is this solution then $y(t) = \sum_{k=0}^{m} b_k x^{(k)}(t)$ is the desired stationary process.

Since white noise is mathematically treacherous we replace (4) by

$$\underline{x}'(t) = \underline{Ax}(t) + \underline{w}_k(t) \tag{4'}$$

where $w_k(t) = \text{col}(0,\ldots,0,u_k(t)/a_n))$ and u_k is a stationary Gaussian (Markov) process with mean 0, covariance $R_k(t) = ke^{-k|t|}$, and (hence) spectral density $f_k(\lambda) = \frac{k^2}{k^2+\lambda^2}$. Clearly as $k\to\infty$ u_k behaves more and more like noise.

$$\text{Let } g(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} P(i\lambda)\Big/Q(i\lambda)e^{it\lambda}\,d\lambda \quad .$$

Because of the special properties of P and Q, $\overline{P(s)} = P(\overline{s})$, $\overline{Q(s)} = Q(\overline{s})$, and the location of the roots of Q, $g(t)$ is easily seen to be real and $g(t) \equiv 0$ if $t<0$. Also $|g(t)| \leq ke^{-st}$ for some $s>0$.

Consider $v_k(t) = \int_0^t g(t-s)u_k(s)ds$. Let

$$r_k(t,t+h) = E(v_k(t)v_k(t+h)) = \int_0^t\int_0^{t+h} g(t+h-\nu_1)g(t-\nu_2)R_k(\nu_1-\nu_2)d\nu_1 d\nu_2.$$

As $t\to\infty$ $\Gamma_k(t,t+h) \to \Gamma_k(h) = \int_{-\infty}^{h}\int_{-\infty}^{0} g(h-\nu_1)g(-\nu_2)R_k(\nu_1-\nu_2)d\nu_1 d\nu_2$

Γ_k has spectral density $f(\lambda)f_k(\lambda)$. Hence asymptotically $v_k(t)$ is stationary with spectral density $f(\lambda)f_k(\lambda)$. As $k \to \infty$ v_k asymptotically is (statistically) just $y(t)$, the desired process.

Since $v_k(t) = \int_0^t g(t-s)u_k(s)ds$, taking Laplace transforms $\hat{v}_k(s) = \int_0^\infty e^{-st}v_k(t)dt = P(s)/Q(s)\ \hat{u}_k(s)$ if $s>0$. ($U_k(t)$ is almost surely transformable because

$$E|\hat{u}_k(s)| \leq \int_0^\infty e^{-st}E|u_k(t)|dt = E(|u_k(0)|)/s.)$$

Hence $(a_0s^{-n}+a_1s^{1-n}+\ldots+a_n)\hat{v}_k(s) = (b_0s^{-n}+\ldots+b_ms^{m-n})\hat{u}_k(s)$

$$a_0v_k^{(-n)}(t) + a_1v_k^{(1-n)}(t) + \ldots + a_nv_k(t)$$

(21)

$$= b_0u_k^{(-n)}(t) + \ldots + b_mu_k^{(m-n)}(t)$$

where

$$v^{-1}(t) = \int_0^t v(s)ds,\ v^{-n-1}(t) = \int_0^t v^{-n}(s)ds,\ \text{etc.}$$

Now consider the stochastic differential equation

$$a_n\xi^{(n)}(t) + \ldots + a_0\xi(t) = u_k(t) \quad .$$

This is equivalent to the matrix equation (4') $\underline{x}'(t) = \underline{Ax}(t)+\underline{w}_k(t)$
If $\underline{\Phi}(t)$ is a solution of the matrix equation $\underline{\Phi}' = \underline{A\Phi}$, $\underline{\Phi}(0) = \underline{I}$ then

$$\text{(22)} \qquad \underline{x}(t) = \underline{\Phi}(t-a)\underline{x}(a) + \int_a^t \underline{\Phi}(t-s)\underline{w}_k(s)ds \quad .$$

Setting a=0, $\underline{x}(a)=0$, it is clear that $x_n(t)$, the last component of $\underline{x}$, solves $a_0x^{(-n)}(t) +\ldots+ a_nx(t) = u_k^{-1}(t)$ if $x_n(0) = \ldots x_{n-m}(0) = 0$. Since (21) has a unique solution with specified initial condition we see that $v_k(t) = \sum_{j=0}^{n} b_jx_{j+1}(t)$ if $\underline{x}(0)$ is appropriately chosen. But since A has only eigenvalues with negative real parts, $|\Phi(t)| \leq ke^{-\sigma t}$ for some $\sigma>0$ and it is easily seen that no matter what $\underline{x}(0)$ is $\underline{x}(t)$ is asymptotically the same. Hence for <u>any</u> initial condition $\sum b_jx_{j+1}(t)$ is asymptotically stationary with density $f(\lambda)f_k(\lambda)$.

Now let $a\to-\infty$ in (6). Clearly $\underline{x}(t)$ is stationary if and only if $x(t) = \int_{-\infty}^t \underline{\Phi}(t-s)\underline{w}_k(s)ds$. Hence (22) has a unique stationary solution. If $\underline{x}$ is stationary $\sum b_jx_{j+1}(t)$ will be stationary and since it tends toward the desired stationary process it must in fact be the desired process. If we let $y_k(t) = \sum_{j=0}^{m} b_jx_{j+1}(t)$ where $\underline{x} = (x_1,\ldots,x_n)$ is the unique stationary with spectrum $f(\lambda)f_k(\lambda)$. Letting $k\to\infty$ we then get the desired result.

BELLCOMM, INC.

APPENDIX 2

KARHUNEN'S METHOD

In this appendix we show how to compute the statistical properties of the solution to a constant coefficient linear ordinary differential equation with a stationary stochastic driving function.

Any stationary process x(t) has a spectral representation $x(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dz_x(\lambda)$ where $z(\lambda)$ is a stochastic process with orthogonal increments. The correlation function $B_x(h) = E(x(h)x(0))$ is also known to be representable in the form $B_x(h) = \int_{-\infty}^{\infty} e^{i\lambda h} dF_x(\lambda)$ where F is some monotone function of bounded variation. F_x and z_x are related by $F_x(\lambda+\Delta\lambda)-F_x(\lambda) = E|z(\lambda+\Gamma\lambda)-z(\lambda)|^2$. It then follows that for any determined z_x-integrable functions f and g

$$E\left\{\int_{-\infty}^{\infty} f(\lambda)dz_x(\lambda) \int_{-\infty}^{\infty} \overline{g(\lambda)dz_x(\lambda)}\right\} = \int_{-\infty}^{\infty} f(\lambda)\overline{g(\lambda)}dF_x(\lambda) \quad .$$

Karhunen's result can now be stated: (11) has a unique stationary solution x(t) which has the representation

$$x(t) = \int_{-\infty}^{\infty} \frac{e^{i\ t} dz_R(\lambda)}{(i\lambda)^n b_{n-1}(i\lambda)^{n-1}+\ldots+b_1 i\lambda+b_0} \quad .$$

($z_R(\lambda)$ is the spectral process for R(t)). Hence

$$\mathrm{cov}(x(t_1)x(t_2)) = \int_{-\infty}^{\infty} \frac{e^{i\lambda t_1}}{(i\lambda)^n+\ldots+b_0} \frac{e^{-i\lambda t_2}}{(-i\lambda)^n+b_{n-1}(-i\lambda)^{n-1}+\ldots+b_0} \frac{dF_R(\lambda)}{}$$

If $R(t)$ has a natural spectral density $f_R(\lambda)$ this becomes

$$\mathrm{cov}(x(t+h)x(t)) = \int_{-\infty}^{\infty} \frac{e^{i\lambda h}}{|(i\lambda)^n+b_{n-1}(i\lambda)^{n-1}+\ldots+b_0|^2} f_R(\lambda)d\lambda \quad .$$

More generally the covariance between derivatives of x can also be computed: if $0 \leq k, \ell \leq n$

$$(23) \quad \mathrm{cov}(x^{(k)}(t+h)x^{(\ell)}(t)) = \int_{-\infty}^{\infty} \frac{(i\lambda)^k(-i\lambda)^\ell e^{i\lambda h}}{|(i\lambda)^n+\ldots+b_0|^2} f_R(\lambda)d\lambda \quad .$$

This is often much more easily computable than is (10), especially in the case h=0.

BELLCOMM, INC.

REFERENCES

1. Skidmore, L. J., Probability and Random Processes, Presented in course at UCLA Extension entitled "Space Navigation and Guidance," October 16, 1967.

2. Coddington, E. A. and Levinson, N., Theory of Ordinary Differential Equations, New York, McGraw-Hill Book Co., Inc., 1955.

3. Karhunen, K., "Linear Transformation Stationaarer Stochastischer Prozesse," Comptes Rendus du duxieme Congress des Mathematicers Scandinaves, Kobenhavn, 1946.

4. Kendall, M. G. and Stuart, A., The Advanced Theory of Statistics, Volume 2, London, Griffin and Co., 1966.